

Introducción a ML y GenAI

Regresión Logística

Ariel Ramos Vela

25-09-2024

Agenda

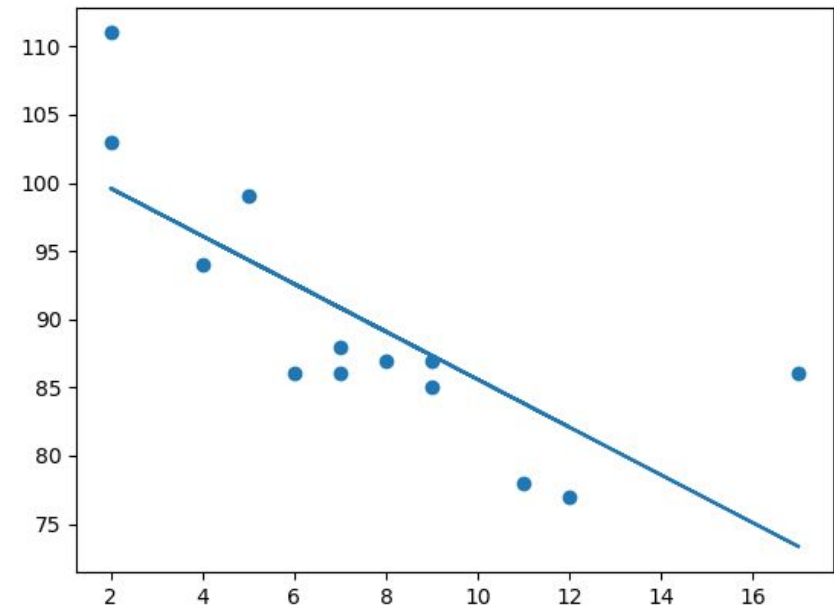
1. Recordatorio: Regresión Lineal
2. ¿Qué es la regresión logística?
3. Función Sigmoide (Sigmoid function)
4. Función de Cost (Cost Function)
5. Gradient Descent
6. Conclusiones
7. Taller 5

Recordatorio - Regresión Lineal

¿Qué es Regresión Lineal?

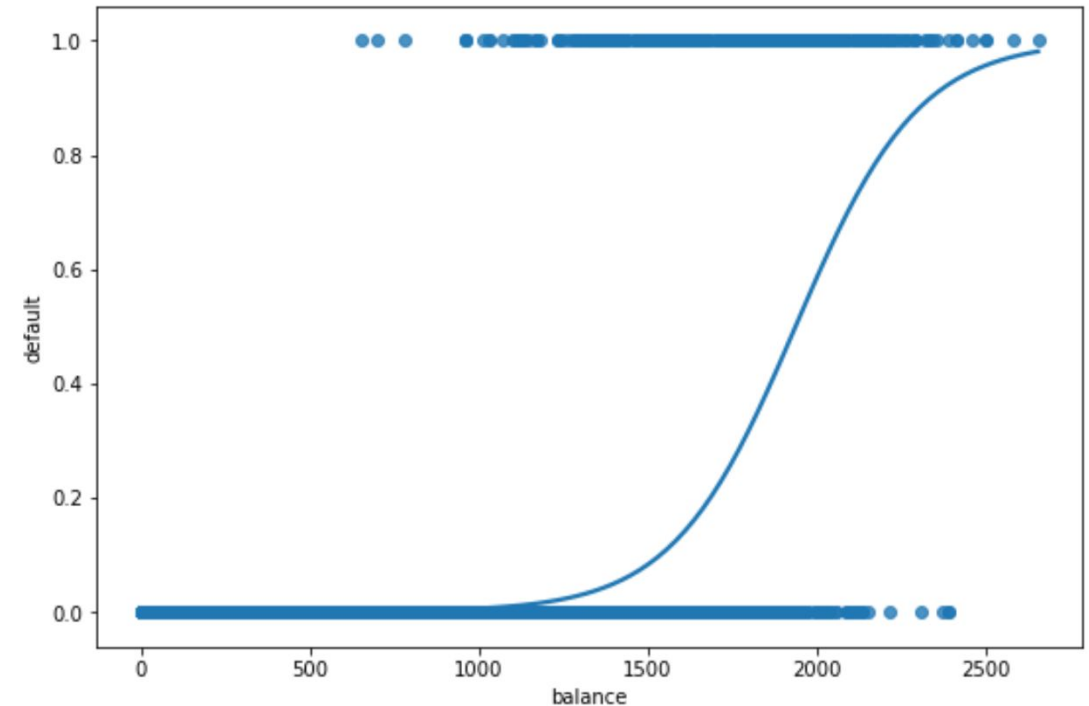
- **Objetivo:** Predecir una variable continua a partir de variables independientes.
- **Ejemplo:** Predecir el precio de una casa en función del tamaño, ubicación, etc.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$



¿Qué es Regresión Logística?

- **Regresión Logística** es una técnica de clasificación utilizada para predecir variables categóricas binarias (0 o 1).
- **Objetivo:** Modelar la probabilidad de que un evento ocurra (por ejemplo, "sobrevivió" o "no sobrevivió").
- **Aplicación:** Ideal para problemas de clasificación binaria (por ejemplo, sí/no, verdadero/falso).
- **Ejemplo:** ¿Sobrevivirá o no una persona en función de su edad, clase, sexo, etc.?



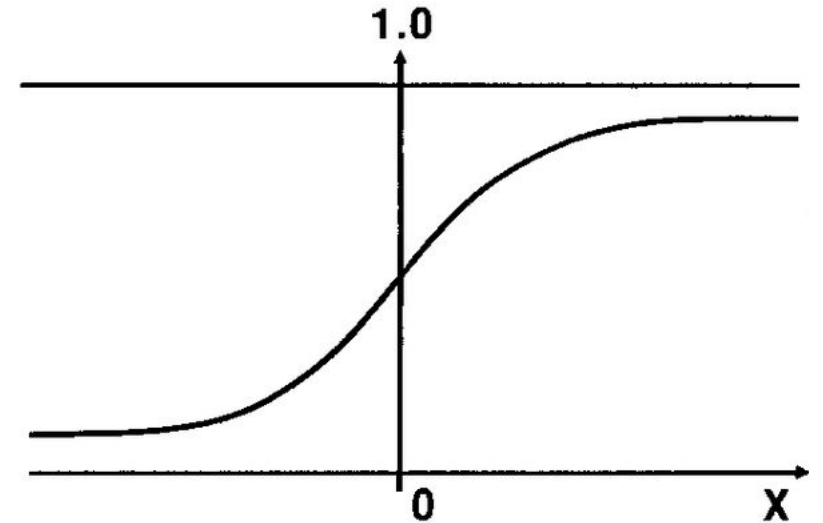
Diferencias entre Regresión Lineal y Logística

Regresión Lineal	Regresión Logística
Predice valores continuos	Predice valores categóricos (binarios)
Usa la función lineal	Usa la función sigmoide
La salida puede ser cualquier valor	La salida es una probabilidad entre 0 y 1

Función Sigmoide

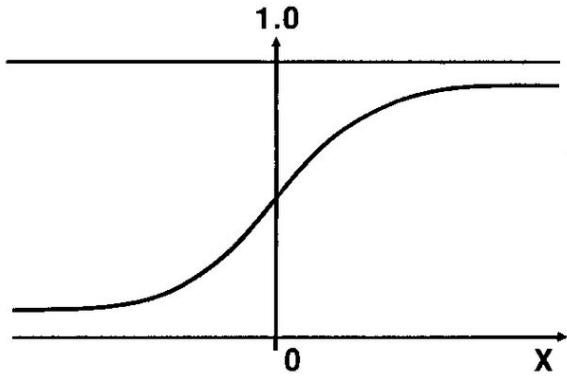
- La **función sigmoide** es crucial en Regresión Logística.
- **Salida:** Un valor entre 0 y 1, interpretado como la probabilidad de que el evento ocurra.
- Convierte la salida de una combinación lineal en una probabilidad.
- **Ejemplo:** Si $\sigma(z)=0.8$, existe un 80% de probabilidad de que el evento ocurra (sobrevivir).

$$\sigma(z) = \frac{1}{1+e^{-z}}$$



¿Cuál es el valor de z ?

$$\sigma(z) = \frac{1}{1+e^{-z}}$$



$$y = mx + b$$

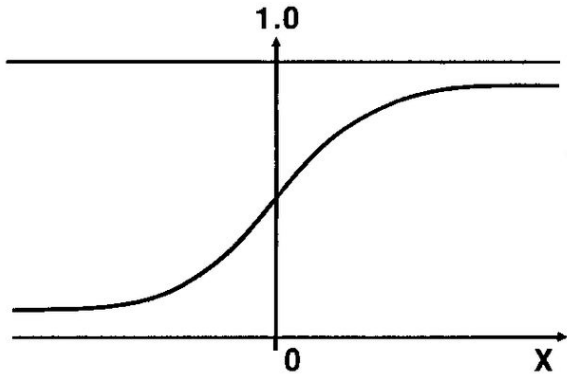
$$\hat{y} = \sigma(mx + b)$$

Ecuación lineal (para 1
feature)

Para regresión
logística

¿Cuál es el valor de z ?

$$\sigma(z) = \frac{1}{1+e^{-z}}$$



$$y = mx + b$$

Ecuación lineal (para 1 feature)

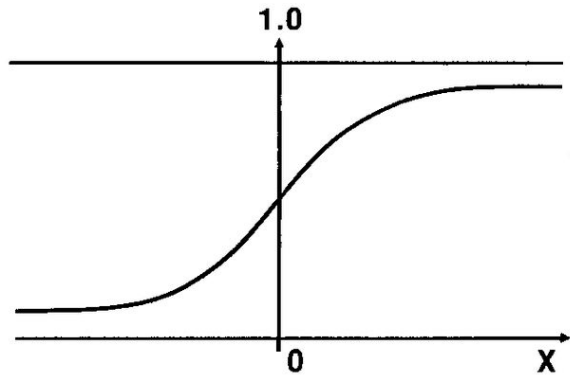
$$\hat{y} = \sigma(mx + b)$$

Para regresión logística

¿Y si tenemos más features?

¿Cuál es el valor de z ?

$$\sigma(z) = \frac{1}{1+e^{-z}}$$



$$y = mx + b$$

Ecuación lineal (para 1 feature)

$$\hat{y} = \sigma(mx + b)$$

Para regresión logística

¿Y si tenemos más features?

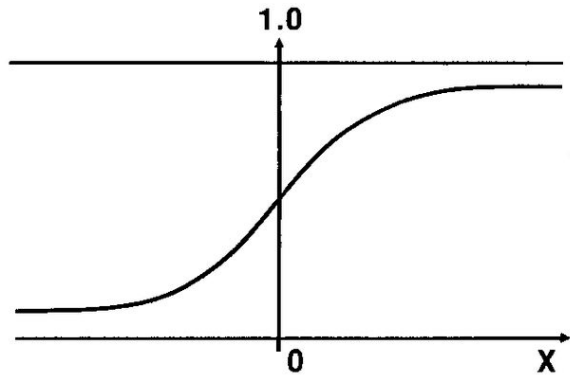
$$\hat{y} = \sigma(W_n x_n + W_{n-1} x_{n-1} + W_{n-2} x_{n-2} + \dots + W_1 x_1 + b)$$

W :son los parámetros que el algoritmo debe aprender (weights)

X : Features

¿Cuál es el valor de z ?

$$\sigma(z) = \frac{1}{1+e^{-z}}$$



$$\hat{y} = mx + b$$

Ecuación lineal (para 1 feature)

$$\hat{y} = \sigma(mx + b)$$

Para regresión logística

¿Y si tenemos más features?

$$\hat{y} = \sigma(W_n x_n + W_{n-1} x_{n-1} + W_{n-2} x_{n-2} + \dots + W_1 x_1 + b)$$

W :son los parámetros que el algoritmo debe aprender (weights)

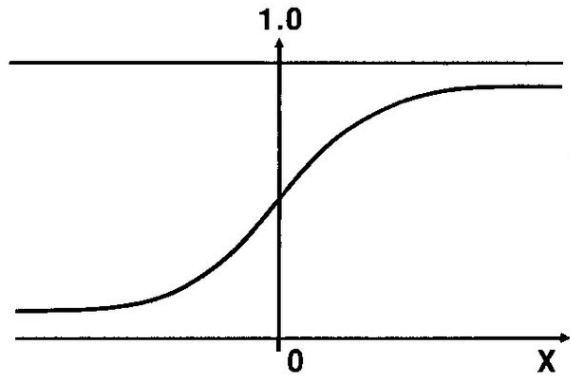
X : Features

$$\hat{Y} = \sigma(XW + b)$$

Matrix form

¿Cuál es el valor de z ?

$$\sigma(z) = \frac{1}{1+e^{-z}}$$



$$\hat{y} = mx + b$$

Ecuación lineal (para 1 feature)

$$\hat{y} = \sigma(mx + b)$$

Para regresión logística

¿Y si tenemos más features?

$$\hat{y} = \sigma(W_n x_n + W_{n-1} x_{n-1} + W_{n-2} x_{n-2} + \dots + W_1 x_1 + b)$$

W :son los parámetros que el algoritmo debe aprender (weights)

X : Features

$$\hat{Y} = \sigma(XW + b)$$

Matrix form

Si $\hat{y} > 0.5$, entonces **Survived (1)**

Si $\hat{y} > 0.5$, entonces **Not Survived (0)**

Función de Coste (Cost Function) para Regresión Logística)

- Recapitulación: para regression lineal $cost = \frac{1}{2m} \sum_{i=1}^m (\hat{y} - y)^2$

Función de Coste (Cost Function) para Regresión Logística

- Recapitulación: para regression lineal $cost = \frac{1}{2m} \sum_{i=1}^m (\hat{y} - y)^2$

- Para regresión logística:

$$cost = \frac{1}{m} \sum_{i=1}^m [y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

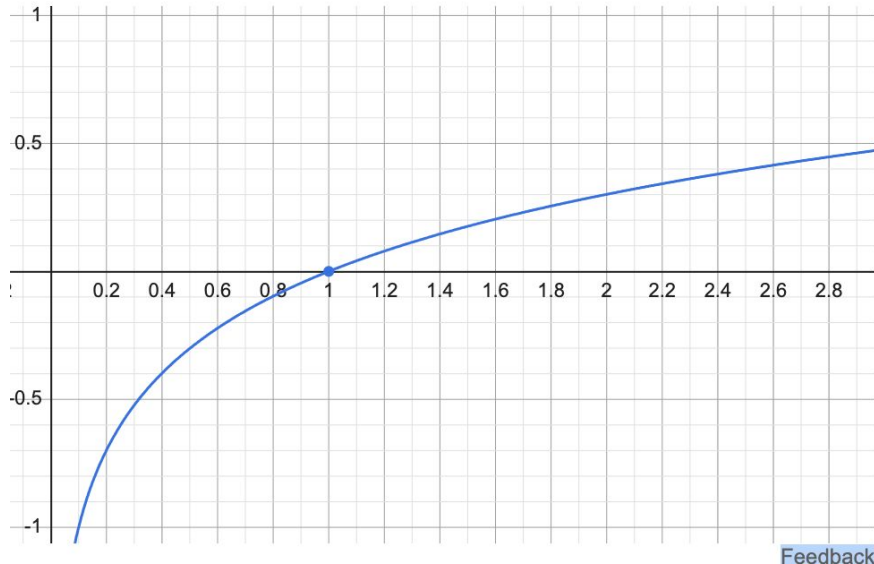
$$\hat{Y} = \sigma(\mathbf{XW} + \mathbf{b})$$

Función de Coste (Cost Function) para Regresión Logística

- Recapitulación: para regression lineal $cost = \frac{1}{2m} \sum_{i=1}^m (\hat{y} - y)^2$
- Para regresión logística:

$$cost = \frac{1}{m} \sum_{i=1}^m [y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

$$\hat{Y} = \sigma(XW + b)$$



Derivada de la Función de coste

$$cost = \frac{1}{m} \sum_{i=1}^m [y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

$$\hat{Y} = \sigma(\mathbf{XW} + b)$$

$$\frac{d(cost)}{d(W)} = \frac{d(cost)}{d(\hat{y})} * \frac{d(\hat{y})}{d(\sigma)} * \frac{d(\sigma)}{dW}$$

$$\frac{d(cost)}{d(b)} = \frac{d(cost)}{d(\hat{y})} * \frac{d(\hat{y})}{d(\sigma)} * \frac{d(\sigma)}{db} =$$

$$\frac{d(cost)}{d(W)} = \frac{1}{m} X(A - Y)$$

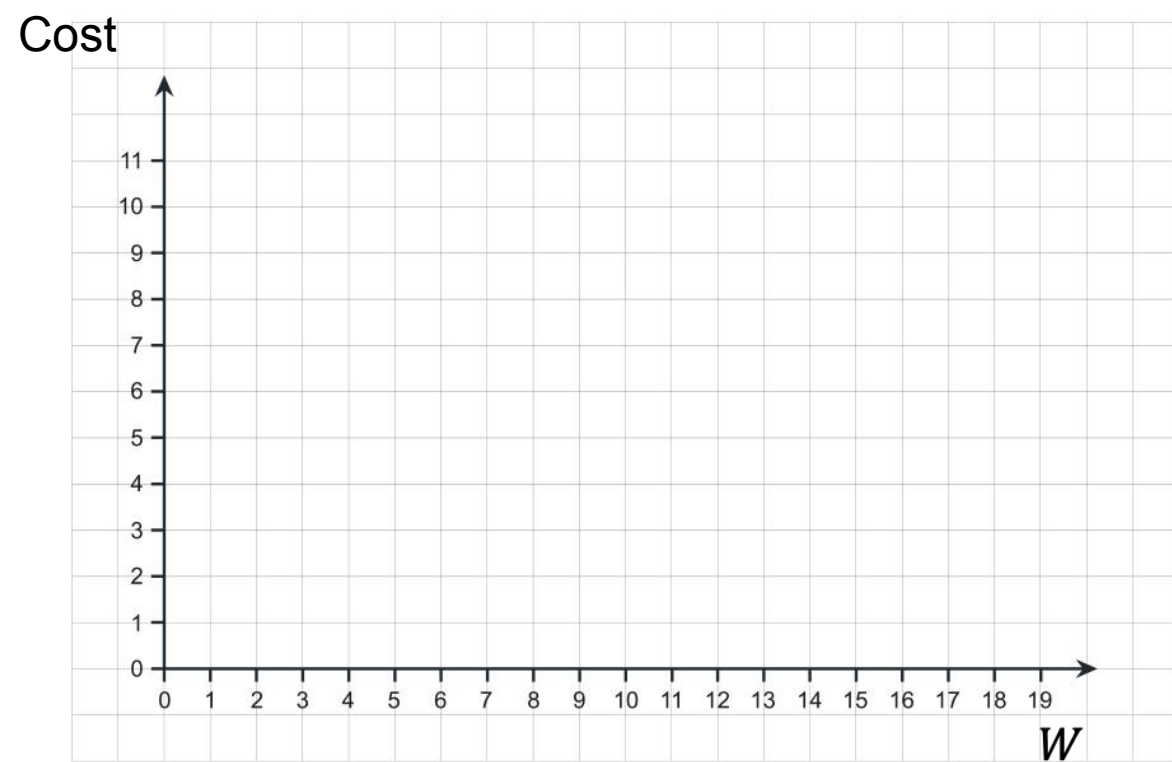
$$\frac{d(cost)}{d(b)} = \frac{1}{m} (A - Y)$$

Donde:

$$A = \sigma(Z)$$

$$Z = \mathbf{XW} + b$$

Algoritmo de Gradient Descent



- Inicializar W y b a cero.
- Bucle (loop) e.g. 1000 times

$$\hat{Y} = \sigma(\mathbf{XW} + b)$$

$$cost = \frac{1}{m} \sum_{i=1}^m [y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

$$\text{Compute } \frac{d(cost)}{dW}$$

$$\text{Compute } \frac{d(cost)}{db}$$

$$W = W - \alpha * \frac{d(cost)}{dW}$$

$$b = -\alpha * \frac{d(cost)}{db}$$

Ventajas y Desventajas de la Regresión Logística

- **Ventajas:**

- Fácil de interpretar.
- La salida (output) son probabilidades .

- **Desventajas:**

- Solo para clasificación binaria.
- Sensible a outliers si no se aplican técnicas adecuadas.

Conclusiones

- La regresión logística es una herramienta poderosa para problemas de clasificación binaria.
- Su fuerza reside en modelar probabilidades y proporcionar interpretaciones claras.
- Con un buen preprocesamiento y ajuste de hiperparámetros, puede ser muy efectiva en múltiples contextos.